63-4-5

THE TECHNOLOGICAL INSTITUTE

NORTHWESTERN UNIVERSITY

EVANSTON ILLINOIS

O. N. R. RESEARCH MEMORANDUM No. 78

The Technological Institute The College of Liberal Arts
Northwestern University

NOTES ON LINEAR INEQUALITIES, I: THE INTERSECTION OF THE NONNEGATIVE ORTHANT WITH COMPLEMENTARY ORTHOGONAL SUBSPACES

by

Adi Ben-Israel*

* Technion-Israel Institute of Technology, Haifa, Israel and Northwestern University, Evanston, Illinois

July, 1963

Acknowledgment: The author is greatly indebted to Professor A. Charnes and to Dr. A. J. Hoffman for their helpful comments on a draft of this paper.

This research was partly supported by the Office of Naval Research, contract Nonr-1228(10), project NR 047-021, and by the National Science Foundation, project G-14102. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government.

SYSTEMS RESEARCH GROUP

A. Charnes, Director

Abstract

the intersections of the nonnegative orthant in E^n with pairs of complementary orthogonal subspaces are investigated. Applications to linear inequalities and linear programming are then made by using Fredholm's alternative theorem.

Introduction

The theory of linear inequalities, the classic reference on which is [17],[1] is closely related to the theory of linear equations.[2] This relation is the subject of the present paper. We show that the main results in the theory of linear inequalities (in finite dimensional vector spaces over arbitrary ordered fields) follow from two basic facts: (a) Theorem 4 below which is an elementary property of the intersections of the nonnegative orthant with pairs of complementary orthogonal subspaces,[3] b) Fredholm's "alternative theorem."[4] Thus new proofs, valid in arbitrary ordered fields, are given for some well-known theorems in [18] [13] [2] and [3] (corollaries 7, 8 and 9 below) and for the duality theorem of linear programming[5] (remark 10 below).

1/ See also [1], [8] and the bibliography in [16]. pp. 305-322.

2/ E.g., [15] and [8], § 16.

3/ Studied earlier in [5] [6] and [7].

4/ E.G., [20], p. 340.

5/ A relation between duality and orthogonality was recently given in [19].

0. Notations

The notations used in this paper are:

$\{x:f(x)\}$	the set of elements x for which $f(x)$ is true
$\{x\}$	the set consisting of x
ϕ	the empty set
ϵ, $\subset$	respectively element and set containment
$\cup$, $\cap$, $-$	respectively set union, intersection and difference
$\mathcal{F}$	an arbitrary ordered field
E^n	the n-dimensional vector space over $\mathcal{F}$
θ	the zero vector in E^n

$C\{f_i : i=1,\ldots,k\} = \{x : x \epsilon E^n, x = \sum_{i=1}^{k} \alpha_i f_i, \alpha_i \epsilon \mathcal{F}, \alpha_i \geqq 0, i=1,\ldots,k\}$

the cone spanned by the vectors $\{f_i : i=1,\ldots,k\}$ in E^n

$\text{int}\, C\{f_i : i=1,\ldots,k\} = \{x : x \epsilon E^n, x = \sum_{i=1}^{k} \alpha_i f_i, \alpha_i \epsilon \mathcal{F}, \alpha_i > 0, i=1,\ldots,k\}$

the interior of $C\{f_i : i=1,\ldots,k\}$

$\text{bdry}\, C\{f_i : i=1,\ldots,k\} = C\{f_i : i=1,\ldots,k\} - \text{int}\, C\{f_i : i=1,\ldots,k\}$

the boundary of $C\{f_i : i=1,\ldots,k\}$

$\{e_i : i=1,\ldots,n\}$ a fixed orthonormal base in E^n

$E_+^n = C\{e_i : i=1,\ldots,n\}$ the nonnegative orthant in E^n

For $x, y \epsilon E^n$ let:

$x \geqq y$ denote $x - y \epsilon E_+^n$

$x \geq y$ denote $x - y \epsilon E_+^n - \{\theta\}$

$x > y$ denote $x - y \epsilon \text{int}\, E_+^n$

For $x, y \epsilon E^n$ let:

$(x, y) = \sum_{i=1}^{n} x_i y_i$ denote the inner product of x, y.

$\|x\| = \sqrt{(x,x)}$ the norm of x

$x \perp y$ denote $(x, y) = 0$

For a subspace L in E^n let:

$x \perp L$ denote $x \perp y$ for all $y \in L$

$L^\perp = \{x: x \in E^n,\ x \perp L\}$, the orthogonal complement of L

dim L : the dimension of L

$x + L = \{y: y \in E^n,\ y = x + \ell,\ \ell \in L\}$ a translate of L

P_L : the perpendicular projection on L , i.e. $P_L = P_L^2 = P_L^T$,

$L = \{x: x \in E^n,\ P_L x = x\}$

For an $m \times n$ matrix A over $\mathcal{F}$ let:

A^T denote the transpose of A

$N(A) = \{x: x \in E^n,\ Ax = \theta\}$ the null space of A

$R(A^T) = \{y: y \in E^n,\ y = A^T v$ for some $v \in E^m\}$ the range space of A^T

1. Theorem: Let P be a perpendicular projection in E^n. Then the following are equivalent:

(i) $Py = \theta$ has no solution $y \geq \theta$.

(ii) $Px = x$ has some solution $x > \theta$.

Proof: (i) $\Longrightarrow$ (ii)

Proof by induction on rank $P = k$, $k=0,1,\ldots,n$. Since both (i) and (ii) are false for $k = 0$, let first P be a perpendicular projection of rank 1. R(P) , the subspace of all solutions of

(1) $Px = x \qquad x \in E^n$ 1/

is of dimension 1 and therefore representable as:

(2) $R(P) = \{x: x = \alpha u,\ \alpha \in \mathcal{F},\ u$ a non-zero solution of (1)$\}$

1/ E.g., [12], § 41, theorem 2.

Suppose now that (ii) is false. This is possible only in two (not mutually exclusive) cases, where $\{u_j\}_{j=1}^n$ are the coordinates of the u in (2):

Case A: $u_i = 0$ for some $1 \leq i \leq n$.

Case B: $u_k u_\ell < 0$ for some $1 \leq k, \ell \leq n$, i.e., some two components of u are of opposite sign.

In each case consider the vector v, given by its coordinates $\{v_j\}_{j=1}^n$ as follows:

Case A: $v_i = 1$, $v_j = 0$ for $j \neq i$

Case B: $v_k = 1$, $v_\ell = -\frac{u_k}{u_\ell}$, $u_j = 0$ for $j \neq k, \ell$.

The vector v satisfies

(3) $$v \geq \theta$$

(4) $$v \perp u .$$

Combining (4) and (2) we conclude that $v \perp R(P)$ and therefore satisfies [1/]

(5) $$Pv = \theta$$

But (3) and (5) imply (i) to be false, and thus (i) $\Longrightarrow$ (ii) is proved for $k = 1$.

<u>Suppose (i) $\Longrightarrow$ (ii) is true for perpendicular projections of rank k, k=1, . . . , n-1.</u>

Let P be a perpendicular projection of rank k+1, written as

(6) $$P = Q_1 + Q_2$$

where Q_1, Q_2 are perpendicular projections of ranks 1, k respectively satisfying

(7) $$Q_1 Q_2 = O = Q_2 Q_1$$ [2/]

1/ E.g., [12], p. 146.

2/ [12], § 76, theorem 1.

Suppose that P satisfied (i). But $Py = \theta$ if and only if $Q_1 y = \theta = Q_2 y$; hence both Q_1, Q_2 satisfy (i). By the induction hypothesis both Q_1, Q_2 satisfy (ii). Thus there exist vectors: $y > \theta$, $z > \theta$ such that

(8) $$Q_1 y = y \quad \text{and} \quad Q_2 z = z \ .$$

The vector $x = y + z$ is $> \theta$, and, because of (6), (7) and (8), satisfies (1). Thus P satisfies (ii) and the proof of (i) $\Longrightarrow$ (ii) is completed.

<u>(ii) $\Longrightarrow$ (i)</u>

If $Px = x > \theta$ then for any $y \geq \theta$:

$$0 < (x, y) = (Px, y) = (x, Py)$$

therefore $Py \neq \theta$. 1/

∎

2. <u>Corollary</u>: Let $L, L^{\perp}$ be complementary orthogonal subspaces in E^n. Then the following are equivalent:

(i) $L \cap E^n_+ = \{\theta\}$

(ii) $L^{\perp} \cap \operatorname{int} E^n_+ \neq \phi$

<u>Proof</u>: Let $P_{L^{\perp}}$ be the projection P in theorem 1. Statements (i) and (ii) are then equivalent to the corresponding statements in theorem 1.

∎

<u>Remark</u>: A stronger result is given in theorem 4 below.

3. <u>Corollary</u>: Let L be a subspace in E^n, of dimension: $\dim L \leqq n-2$, such that $L \cap E^n_+ = \{\theta\}$. Then L is contained in a subspace M, with $\dim M = \dim L + 1$ and $M \cap E^n_+ = \{\theta\}$.

1/ Only the symmetry of P was used here.

Proof: Let $\dim L = k$, $0 \leq k \leq n-2$; and let $x_1, \ldots, x_{n-k}$ be an orthonormal basis of $L^{\perp}$ with $x_1 \in L^{\perp} \cap \operatorname{int} E_+^n$ (by corollary 2 this is always possible). The subspace $M = \{L + \lambda x_2 : \lambda \in \mathcal{F}\}$ contains L and is of dimension: $\dim M = \dim L + 1$. Since $M \perp x_1$ and $x_1 \in \operatorname{int} E_+^n$ it follows that $M \cap E_+^n = \{\theta\}$ ▨

Remark: We have shown that any subspace L of E^n, with dimension $\leqq n-2$, satisfying $L \cap E_+^n = \{\theta\}$ can be extended to a subspace M, $\dim M = \dim L + 1$, and the same property. The maximal subspace $H \supset L$ with $H \cap E_+^n = \{\theta\}$ is a hyperplane.

4. Theorem: Let L be a subspace of E^n. Then the following are equivalent:

(i) $L \cap E_+^n = \{\theta\}$

(ii) $L^{\perp}$ has a basis in $\operatorname{int} E_+^n$

(iii) For every $x \in E^n$, $\{x + L\} \cap E_+^n$ is bounded, maybe empty.

Proof: The part (ii) $\Longrightarrow$ (i) follows from corollary 2. Also (iii) $\Longrightarrow$ (i) is obvious, for if (i) is false then (iii) is false with $x = \theta$. It remains to prove that (i) $\Longrightarrow$ (ii) and (i) $\Longrightarrow$ (iii).

(i) $\Longrightarrow$ (ii)

Proof by induction on $\dim L = n-k$, $k = 0, 1, \ldots, n$. The case $k=0$ is trivial and the case $k=1$ is proved exactly as in the part: (i) $\Longrightarrow$ (ii) of theorem 1. Suppose that (i) $\Longrightarrow$ (ii) was proved for subspaces of dimension: $n-k$, and let L be a subspace of dimension: $\dim L = n-k-1$, such that $L \cap E_+^n = \{\theta\}$.

By corollary 3 $L \subset M$ with $\dim M = n-k$ and $M \cap E_+^n = \{\theta\}$. By the induction hypothesis $M^\perp$ has a basis in $\operatorname{int} E_+^n$, say $\{g_i : i=1,\dots,k\}$.

Let $\{f_i : i=1,\dots,n-k-1\}$ be an orthonormal basis of L and $\{f_i : i=1,\dots,n-k\}$ be an orthonormal basis of M. Consider the vector

$$g_{k+1} = f_{n-k} + \lambda g_1$$

where $\lambda \in \mathcal{F}$ satisfied:

$$\lambda > \max_{f_{n-k,i} < 0} \frac{|f_{n-k,i}|}{g_{1,i}}$$

The vector g_{k+1} is: (a) in $\operatorname{int} E_+^n$, (b) orthogonal to $f_1,\dots,f_{n-k-1}$ and thus in $L^\perp$, (c) linearly independent of $g_1,\dots,g_k$. Therefore $\{g_1,\dots,g_{k+1}\}$ is a basis of $L^\perp$ in $\operatorname{int} E_+^n$.

<u>(i) ==> (iii)</u>

Clearly it is enough to consider vectors $x \in L^\perp$. Let u be any vector in L with $\|u\| = 1$. Since $L \cap E_+^n = \{\theta\}$ it follows that u must have nonzero coordinates of opposite signs. For any $x \in L^\perp$, $\alpha \in \mathcal{F}$, the vector $x + \alpha u$ is $\geq \theta$ only if the scalar α is bounded by

$$(9) \qquad \max\left\{ \max_{\substack{x_i \geq 0 \\ u_i > 0}} \left\{-\frac{x_i}{u_i}\right\},\ \max_{\substack{x_i < 0 \\ u_i > 0}} \left\{\frac{|x_i|}{u_i}\right\} \right\} \leq \alpha \leq \min\left\{ \min_{\substack{x_i \geq 0 \\ u_i < 0}} \left\{\frac{x_i}{|u_i|}\right\},\ \min_{\substack{x_i < 0 \\ u_i < 0}} \left\{-\frac{x_i}{u_i}\right\} \right\}$$

where x_i, u_i are the coordinates of x, u respectively. If the left hand side in (9) exceeds the right hand side then $\{x+L\} \cap E_+^n = \phi$. Otherwise $x + \alpha u \in \{x+L\} \cap E_+^n$ only if $\|x+\alpha u\|^2 = (x,x) + \alpha^2 < \infty$, by (9), which completes the proof.

5. <u>Corollary</u>: Let L be a subspace of E^n of dimension k, $k=1,\dots,n$. Then the following are equivalent:

(i) $L \cap \operatorname{bdry} E^n_+ = C\{e_1,\dots,e_p\}$, $1 \le p \le k$.

(ii) $L^{\perp}$ has a basis in $\operatorname{int} C\{e_{p+1},\dots,e_n\}$.

<u>Proof:</u>

<u>(i) => (ii)</u>

Consider E^p, the space spanned by $\{e_1,\dots,e_p\}$ as a subspace of E^n, and the quotient space of E^n modulo E^p: E^n/E^p. From (i) it follows that $E^p \subset L$. Hence in E^n/E^p the subspace L/E^p satisfies (i) of theorem 4. Therefore $L^{\perp}/E^p$ has a basis in $\operatorname{int}(E^n_+/E^p) = \operatorname{int} C\{e_{p+1},\dots,e_n\}$. But $L^{\perp}/E^p = L^{\perp}$ and (ii) is established.

<u>(ii) => (i)</u>

From (ii) it follows that $C\{e_1,\dots,e_p\} \subset L$ and consequently (i). ▨

<u>Remarks:</u>

(a) If $p = k$ then (ii) can be rewritten as

(ii') $$L^{\perp} \cap E^n_+ = C\{e_{p+1},\dots,e_n\}$$

(b) If $\dim L \leqq n-2$ it can be shown as in corollary 3 that L is contained in a subspace M with $\dim M = \dim L + 1$ and $M \cap \operatorname{bdry} E^n_+ = C\{e_1,\dots,e_p\}$. The maximal subspace H with $H \supset L$ and $H \cap \operatorname{bdry} E^n_+ = C\{e_1,\dots,e_p\}$ is a hyperplane, e.g., [10], p. 316, theorem 33 (2).

6. <u>Corollary</u>: Let $L, L^{\perp}$ be any pair of complementary orthogonal subspaces in E^n. Then there is a vector x in $\operatorname{int} E^n_+$ such that $x = y + z$, $y \in L \cap E^n_+$, $z \in L^{\perp} \cap E^n_+$.

Proof:

Since the case: $L = E^n$, $L^{\perp} = \{\theta\}$ is trivial, let $\dim L = 1,\ldots,n-1$. Now there are three mutually exclusive cases:

(i) $L \cap E^n_+ = \{\theta\}$

(ii) $L \cap \text{bdry}\, E^n_+ = C\{e_1,\ldots,e_p\}$, $1 \leq p \leq \dim L$ and $L \cap \text{int}\, E^n_+ = \phi$

(iii) $L \cap \text{int}\, E^n_+ \neq \phi$.

In case (i) we use corollary 2 and choose x as a vector in $L^{\perp} \cap \text{int}\, E^n_+$. Thus $x = z$ and $y = \theta$.

In case (ii), corollary 5 is used to construct x as $x = y + z$ where $z \in L^{\perp} \cap \text{int}\{e_{p+1},\ldots,e_n\}_+$ and y is any vector in $\text{int}\, C\{e_1,\ldots,e_p\}$. By remark (a) following corollary 5, if $p = \dim L$ then any vector x in $\text{int}\, E^n_+$ can be so represented.

Case (iii) is, by corollary 2, case (i) with $L, L^{\perp}$ permuted.

▨

7. Corollary: (Tucker [18]) Let A be any $m \times n$ matrix over $\mathcal{F}$. Then the following system of equations and inequalities

$$Ax = \theta \qquad\qquad A^T u \geqq \theta$$
$$x \geqq \theta$$

has solutions x^o, u^o satisfying

$$A^T u^o + x^o > \theta .$$

Proof: Follows immediately from corollary 6, by letting $L = R(A^T)$ and using Fredholm's alternative theorem ([20], p. 340) which for a linear operator $A: E^n \to E^m$ can be stated as: R(A^T) and N(A) are complementary orthogonal subspaces in E^n, e.g., [12], § 49.

∎

Remark: This corollary is fundamental in the theories of linear inequalities and linear programming, e.g., [18]. We proved it here as a consequence of two facts:

(a) Corollary 6 which states a simple property of the intersections of E^n_+ with arbitrary pairs of complementary orthoronal subspaces.

(b) Fredholm's alternative theorem

$$E^n = R(A^T) \oplus N(A)$$

which is basic to the theory of linear equations.

The abundance of theorems (e.g., the Minkowski-Farkas-Weyl theorems [2] and their consequences) which follow corollary 7, e.g., [18] emphasizes the merits of a unified treatment of linear inequalities and equations, e.g., [15].

8. Corollary (Jackson [13]. Charnes-Cooper [2]): Let A be any $m \times n$ matrix over $\mathcal{F}$. Then the following are equivalent:

(i) $Ax = \theta$ has no solution $x \geq \theta$.

(ii) $A^T u > \theta$ has solutions.

(iii) For any $b \in E^m$ the set $\{x : x \in E^n, Ax = b, x \geqq \theta\}$ is bounded, maybe empty.

Proof: Setting, by Fredholm's alternative theorem, $L = N(A)$ and $L^{\perp} = R(A^T)$ it follows that statements (i), (ii) and (iii) are equivalent to the corresponding statements in theorem 4.

∎

Remark: In the real case this theorem was proved by Jackson [13]. The part (i) ==> (iii) is close to the "opposite sign theorem" of Charnes-Cooper [2], [3], which states, more precisely, that (i) above is equivalent to:

(iv) The set

$$\{x : x \in E^n,\ Ax = b,\ x \geqq \theta\}$$

is spanned by its extreme points.

Property (iv), see [4], is not restricted to bounded sets when x is in an infinite dimensional vector space.

9. Corollary (Tucker [18]): Let K be a skew symmetric matrix over $\mathcal{F}$.[1/] Then the system of inequalities

$$Kw \geqq \theta \qquad w \geqq \theta$$

has a solution w^o such that

(10) $$Kw^o + w^o > \theta \ .$$

Proof: Consider the system of equations and inequalities

$$(I,\ K^T)\begin{pmatrix} x \\ y \end{pmatrix} = \theta \qquad \begin{pmatrix} I \\ K \end{pmatrix} u \geqq \theta$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \geqq \theta$$

by corollary 7 this system has solutions $\begin{pmatrix} x^o \\ y^o \end{pmatrix}$ and u^o such that

(11) $$\begin{pmatrix} x^o \\ y^o \end{pmatrix} + \begin{pmatrix} I \\ K \end{pmatrix} u^o > \theta \ .$$

Combining (11) and the fact that $x^o = -K^T y^o = Ky^o$, it follows that $w^o = y^o + u^o$ satisfies (10). ∎

1/ I.e., $K = -K^T$.

10. <u>Remark</u>: Corollary 9 was used to prove the duality theorem of linear programming, e.g., [9] and [11]. We conclude this paper by outlining an alternative proof which, like our other results above, rests upon the "Fredholm alternative" theorem.

Consider the pair of dual problems

$$\begin{array}{lll} \text{maximize} & c^T x \qquad\qquad & \text{minimize} \quad w^T b \\ & Ax \leqq b & \qquad\qquad w^T A \geqq c^T \\ & x \geqq \theta & \qquad\qquad w \geqq \theta \end{array}$$

where A is an $m \times n$ matrix over $\mathcal{F}$, $b \in E^m$ and $c \in E^n$.

To these problems there corresponds the $(m+1) \times (m+n+1)$ matrix:

$$B_t = \left(\begin{array}{c|c|c} t & -c^T & \theta^T \\ \hline -b & A & I \end{array}\right)$$

where t is in $\mathcal{F}$. For any given value of t we consider the subspaces $N(B_t)$ and $R(B_t^T)$ --which are complementary orthogonal by Fredholm's theorem--and their intersections with E_+^{m+n+1}:

$$N(B_t) \cap E_+^{m+n+1} = \left\{ \begin{pmatrix} \alpha \\ x \\ y \end{pmatrix} : \begin{array}{r} \alpha t - c^T x = 0 \\ -\alpha b + Ax + y = \theta \\ \alpha \geqq 0, \ x, y \geqq \theta \end{array} \right\}$$

$$R(B_t^T) \cap E_+^{m+n+1} = \left\{ \begin{pmatrix} \beta t - b^T w \\ -\beta c + A^T w \\ w \end{pmatrix} : \begin{array}{r} \beta t - w^T b \geqq 0 \\ -\beta c^T + w^T A \geqq \theta^T \\ w \geqq \theta \end{array} \right\} .$$

The duality theorem of linear programming (as well as the "complementary slackness" property, [11], which is the statement that $R(B_t^T) \cap E_+^{m+n+1}$ and $N(B_t) \cap E_+^{m+n+1}$ are orthogonal sets) 1/ follows now by considering the above intersections; the keys to the whole situation being the vanishing of the scalars α, β and the value of t. The details are left to the reader.

1/ A similar relation between duality and orthogonality was studied by Tucker in [19].

REFERENCES

[1] S. N. Cernikov: Uspehi Mat. Nauk N.S. 8, No. 2(54), 1953, pp. 7-73.

[2] A. Charnes and W. W. Cooper: Proc. Nat. Acad. Sci. U.S.A., 44, 1958, pp. 914-916.

[3] A. Charnes and W. W. Cooper: Management Models and Industrial Applications of Linear Programming, J. Wiley and Sons, New York, 1961.

[4] A. Charnes, W. W. Cooper and K. Kortanek: Proc. Nat. Acad. Sci. U.S.A. 48, 1962, pp. 783-786.

[5] C. Davis: Michigan Math. J. 1, 1952, pp. 163-168.

[6] C. Davis: Michigan Math. J. 2, 1953, pp. 23-25.

[7] C. Davis: Amer. J. Math. 76, 1954, pp. 733-746.

[8] Ky Fan: pp. 99-156 in [16].

[9] D. Gale, H. W. Kuhn and A. W. Tucker: pp. 317-319 in [14].

[10] M. Gerstenhaber: pp. 298-316 in [14].

[11] A. J. Goldman and A. W. Tucker: pp. 53-97 in [16].

[12] P. R. Halmos: Finite Dimensional Vector Spaces, D. Van Nostard Co., 2nd edition, 1958.

[13] J. R. Jackson: Pacific J. Math. 4, 1954, pp. 29-36.

[14] T. C. Koopmans (editor): Activity Analysis of Production and Allocation, Cowles Monograph No. 13, J. Wiley and Sons, New York, 1951.

[15] H. W. Kuhn: Amer. Math. Monthly 63, 1956, pp. 217-232.

[16] H. W. Kuhn and A. W. Tucker (editors): "Linear Inequalities and Related Systems," Annals of Mathematics Studies, No. 38, Princeton University Press, Princeton, 1956.

[17] T. S. Motzkin: "Beiträge sur Theorie der linearen Ungleichungen," Inaugural Dissertation, Basel, 1933, Jerusalem, Azriel, 1936.

[18] A. W. Tucker: pp. 3-18 in [16].

[19] A. W. Tucker: "Simplex Method and Theory, Notes on Linear Programming and Extensions, part 62," The Rand Corporation, Santa Monica, Memorandum RM-3199-PR, June 1962.

[20] A. C. Zaanen: Linear Analysis, North-Holland Publishing Co., 1953.